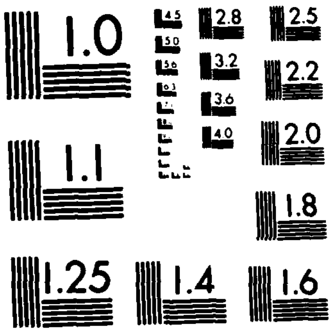
AD-A145 569 NONPARAMETRIC MAXIMUM PENALIZED LIKELIHOOD ETTIMATION 1/1
OF A DENSITY FROM A..(U) SOUTH CAROLINA UNIV COLUMBIA
DEPT OF MATHEMATICS AND STATISTI.. A M LUBECKE ET AL.
UNCLASSIFIED JUN 84 TR-96 AFOSR-TR-84-0668 AFOSR-84-0156 F/G 12/1 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

Department of Mathematics
and Statistics
The University of South Carolina
Columbia, South Carolina 29208

(4)

AD-A145 569

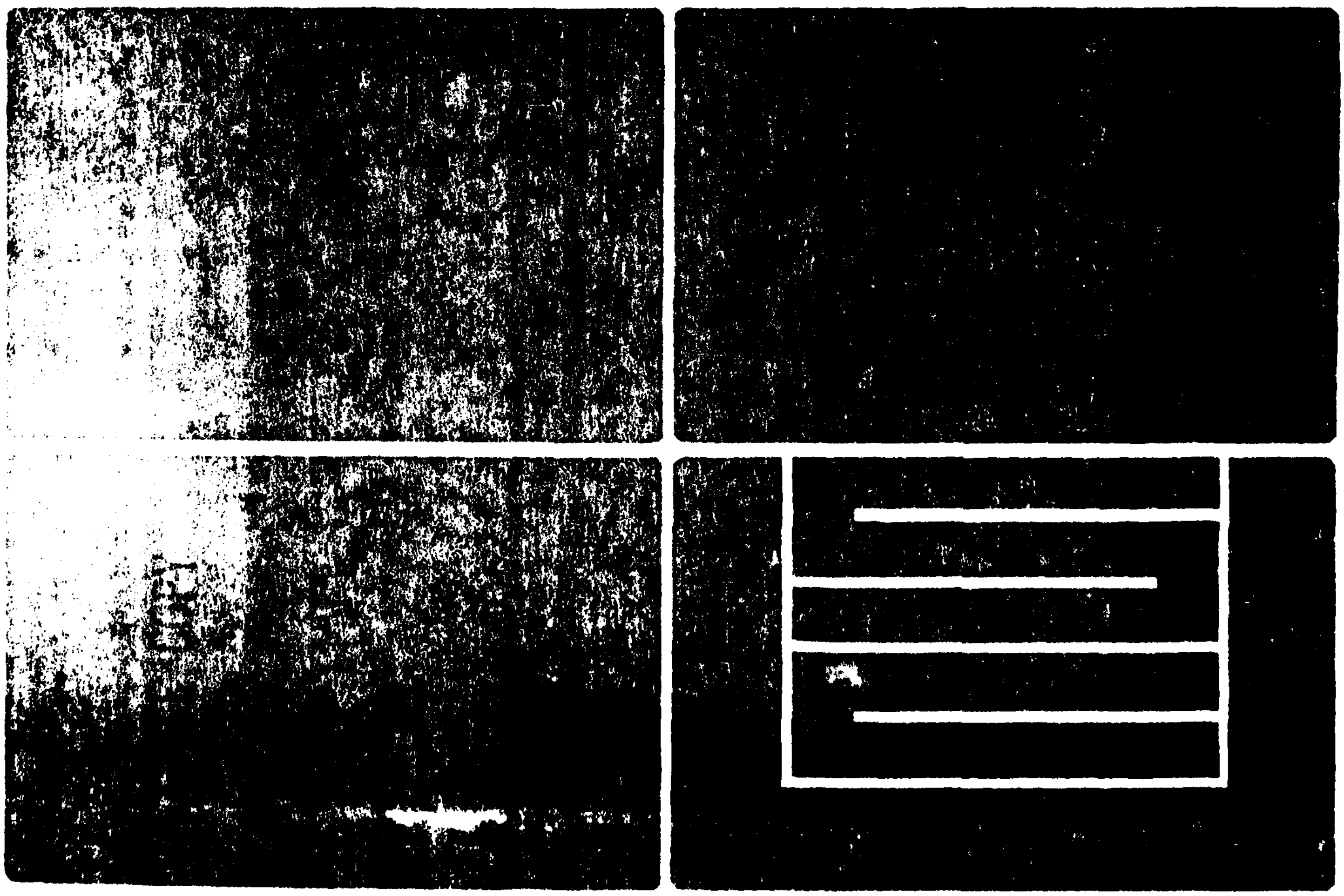
NONPARAMETRIC MAXIMUM PENALIZED LIKELIHOOD ESTIMATION OF A DENSITY FROM ARBITRARILY RIGHT-CENSORED OBSERVATIONS *

by

André M. Lubecke and W. J. Padgett

University of South Carolina

Statistics Technical Report No. 96
62G05-9



NONPARAMETRIC MAXIMUM PENALIZED LIKELIHOOD
ESTIMATION OF A DENSITY FROM ARBITRARILY
RIGHT-CENSORED OBSERVATIONS *

by

André M. Lubecke and W. J. Padgett

University of South Carolina
Statistics Technical Report No. 96
62G05-9

June, 1984

Department of Mathematics and Statistics
University of South Carolina
Columbia, South Carolina 29208

* Research supported by the United States Air Force Office of Scientific Research and Army Research Office under Grant No. AFOSR-84-0156.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 84-0668
6a NAME OF PERFORMING ORGANIZATION University of South Carolina 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c ADDRESS (City, State, and ZIP Code) Department of Mathematics & Statistics Columbia SC 29208	7b ADDRESS (City, State, and ZIP Code) Directorate of Mathematical & Information Sciences, AFOSR, Bolling AFB DC 20332
8a NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b OFFICE SYMBOL (If applicable) NM	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-84-0156
8c ADDRESS (City, State, and ZIP Code) Bolling AFB DC 20332	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61102F PROJECT NO. 2304 TASK NO. A5 WORK UNIT ACCESSION NO.

11 TITLE (Include Security Classification)
NONPARAMETRIC MAXIMUM PENALIZED LIKELIHOOD ESTIMATION OF A DENSITY FROM ARBITRARILY RIGHT-CENSORED OBSERVATIONS.

12 PERSONAL AUTHOR(S)
Andre M. Lubecke and W.J. Padgett

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM ______ TO ______	26 JUN 84	17

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Maximum likelihood estimation; existence and uniqueness; Good and Gaskins' first MPLE; survival estimation; random censorship; nonparametric density estimation; reliability.

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Based on arbitrarily right-censored observations from a probability density function f^o, the existence and uniqueness of the maximum penalized likelihood estimator (MPLE) of f^o is proven. In particular, the "first MPLE of Good and Gaskins" of a density defined on $[0,\infty)$ is shown to exist and to be unique under arbitrary right-censorship. Furthermore, the MPLE is in the form of an exponential spline with knots at the observed censored and uncensored data points.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL CPT Brian W. Woodruff	22b TELEPHONE (Include Area Code) (202) 767-5027 22c OFFICE SYMBOL NM

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

84 08 17 102

NONPARAMETRIC MAXIMUM PENALIZED LIKELIHOOD ESTIMATION OF A DENSITY FROM ARBITRARILY RIGHT-CENSORED OBSERVATIONS

André M. Lubecke and W. J. Padgett

Department of Mathematics and Statistics
University of South Carolina
Columbia, South Carolina 29208

Key Words and Phrases: Maximum likelihood estimation; Existence and uniqueness; Good and Gaskins' first MPLE; Survival estimation; Random censorship; Nonparametric density estimation; Reliability.

ABSTRACT

Based on arbitrarily right-censored observations from a probability density function f^0, the existence and uniqueness of the maximum penalized likelihood estimator (MPLE) of f^0 is proven. In particular, the "first MPLE of Good and Gaskins" of a density defined on $[0,\infty)$ is shown to exist and to be unique under arbitrary right-censorship. Furthermore, the MPLE is in the form of an exponential spline with knots at the observed censored and uncensored data points.

1. INTRODUCTION

The problem of nonparametric probability density estimation has been studied for many years. Summaries of results for complete (uncensored) random samples have been listed by Tapia and Thompson (1978), Wertz and Schneider (1979), and Bean and Tsokos (1980), for example. Also, a review of results for censored

A1

samples has been given by Padgett and McNichols (1984). In addition to its importance in theoretical statistics, nonparametric density estimation has been used in hazard analysis, life testing, and reliability, as well as in the areas of nonparametric discrimination and high energy physics (Good and Gaskins, 1971).

One approach to estimating a density function nonparametrically is that of maximum likelihood. Nonparametric maximum likelihood estimates of a probability density function do not exist in general. That is, the likelihood function for a complete sample is unbounded over the class of all possible densities. However, by suitably restricting the class of densities, a nonparametric maximum likelihood estimator (MLE) may be found within the restricted class. For complete samples, the maximum likelihood estimator of a density g was given by Barlow, Bartholomew, Bremner and Brunk (1972) if g was assumed to be either decreasing (nonincreasing) or unimodal with known mode. Wegman (1970a,b) assumed unimodality with unknown mode and found the MLE of the density and studied its properties for complete samples. McNichols and Padgett (1982) studied the nonparametric MLE of monotonic or unimodal densities based on arbitrarily right-censored observations. Even within the class of decreasing (or unimodal) density functions, however, when the largest observation was censored, McNichols and Padgett (1982) had to restrict their estimator to a finite interval $[0,T]$ where T was an arbitrarily large positive number, greater than the largest observation.

Another approach to the problem of nonparametric maximum likelihood estimation of a density from complete samples was proposed by Good and Gaskins (1971). This method allowed any smooth integrable function on the interval of interest (a,b) (which may be finite or infinite) as a possible estimator, but added a "penalty function" to the likelihood. The penalty function penalized a density for its lack of smoothness, so that a very "rough" density would have a smaller likelihood than a "smooth" density, and hence, would not be admissible. De Montricher, Tapia, and

Thompson (1975) showed that the natural mathematical setting for the solution of the maximum penalized likelihood estimation (MPLE) problem of Good and Gaskins (1971) was provided by the Sobolev subspaces of the Hilbert space $L_2(R)$, the square-integrable functions on the real line R. They proved existence and uniqueness results for the MPLE. Later, Klonias (1982) obtained the strong consistency of the MPLE of the density function in appropriate norms. He also derived the "first MPLE of Good and Gaskins" for the case that the density g has support only on the half line, essentially by reflecting g around zero and using results for g having support R.

In this paper we obtain existence and uniqueness results for the nonparametric MPLE of a density g based on arbitrarily right-censored observations from g. General results are first obtained for densities with support $\Omega \subset R$ and penalty function ϕ and then the problem of "Good and Gaskins' first MPLE" is considered for arbitrarily right-censored data observed on R. The existence and uniqueness results are then obtained for densities g with only positive support by using a symmetry argument, reflecting g about zero, and then utilizing the general results for support R. It is also shown that the MPLE is an exponential spline with knots at the data points.

2. NOTATION AND BASIC DEFINITIONS

Let $\Omega \subset R$ be a finite or infinite interval and let f^o denote a probability density function with support in Ω. Let $X_1^0,\ldots,X_n^0$ be n independent identically distributed random variables with common density f^o. Later, X_i^0, $i=1,\ldots,n$, will be interpreted as the true survival times of n items or individuals under observation, where f^o will have support in $[0,\infty)$. Suppose that $U_1,U_2,\ldots,U_n$ is a sequence of constants or random variables which "censor" X_i^0, $i=1,\ldots,n$, on the right. In survival analysis or reliability studies, the U_i's represent possible "loss" times of items or individuals from the test.

The observed data are denoted by the pairs (X_i,Δ_i), $i=1,\ldots,n$, where

$$X_i = \min\{X_i^0,U_i\},\quad \Delta_i = \begin{cases} 1 & \text{if } X_i^0 \le U_i \\ 0 & \text{if } X_i^0 > U_i . \end{cases}$$

It is desired to obtain the MPLE of f^o based on these observations.

In reliability or survival analysis, where f^o has support in $[0,\infty)$, the nature of the censoring depends on the U_i's. (i) If $U_1,\ldots,U_n$ are fixed constants, the observations are time truncated. If all U_i's are equal to the same constant, then the case of Type I censoring results. (ii) If all $U_i = X_{(r)}^0$, the rth order statistic of $X_1^0,\ldots,X_n^0$, then the situation is that of Type II censoring. (iii) If $U_1,\ldots,U_n$ constitute a random sample from a distribution H (usually unknown) and are independent of $X_1^0,\ldots,X_n^0$, then (X_i,Δ_i), $i=1,\ldots,n$, is called a randomly censored sample. See Gill (1980, Ch. 3 and Ex. 4.1.1) for further discussion. An observed value of (X_i,Δ_i) will be denoted by (x_i,d_i).

By $L^p(\Omega)$ we will mean the space of functions v such that $\int_\Omega |v(t)|^p dt < \infty$ with norm $\|v\|_p = [\int_\Omega |v(t)|^p dt]^{1/p}$ for $p \ge 1$. Let $H(\Omega)$ be a manifold in $L^1(\Omega)$.

Following notation similar to that of De Montricher, Tapia, and Thompson (1975), let ϕ denote a functional $\phi\colon H(\Omega) \to R$. Given the arbitrarily right-censored sample (x_i,d_i), $i=1,2,\ldots,n$, the ϕ-penalized likelihood of $v \in H(\Omega)$ is defined by

$$L(v) = \prod_{i=1}^{n} [v(x_i)]^{d_i} [1-V(x_i)]^{1-d_i} \exp(-\phi(v)),$$

where $V(x_i) = \int_{-\infty}^{x_i} v(t)dt$ denotes the cumulative distribution function with density v and ϕ is the penalty function. By the maximum penalized likelihood estimator (MPLE) of f^o corresponding to manifold $H(\Omega)$ and penalty function ϕ, we will mean any solution to the problem:

$$\begin{aligned}&\text{maximize } L(v) \text{ subject to} \\ &v \in H(\Omega),\ \int_\Omega v(t)dt = 1, \text{ and} \\ &v(t) \ge 0 \text{ for all } t \in \Omega.\end{aligned} \tag{2.1}$$

The function $L(v)$ is the censored form of the penalized likelihood of Good and Gaskins (1971).

When $H(\Omega)$ is a Hilbert space, a natural penalty function to use is $\phi(v) = \|v\|^2$, where $\|\cdot\|$ is the norm on $H(\Omega)$. If no reference is given to ϕ when we are considering the MPLE corresponding to a Hilbert space $H(\Omega)$, it is assumed that ϕ is the square of the norm on $H(\Omega)$. A Hilbert space inner product will be denoted by $<\cdot,\cdot>$ so that $<v,v> = \|v\|^2$. When $H(\Omega)$ is a Hilbert space, it is a reproducing kernel Hilbert space (RKHS) if point evaluation is a continuous operation, that is, $v_n \to v$ in $H(\Omega)$ implies that $v_n(t) \to v(t)$ for all $t \in \Omega$. See Goffman and Pedrick (1965) for further details.

3. EXISTENCE AND UNIQUENESS OF AN MPLE

In this section we establish the existence and uniqueness of a solution to problem (2.1) when $H(\Omega)$ is a RKHS. The inner product on $H(\Omega)$ is defined by $<u,v> = \int_\Omega u(t)v(t)dt$ for $u,v \in H(\Omega)$.

Theorem 3.1. Assume that $H(\Omega)$ is a RKHS, integration over Ω is a continuous functional, and D is a closed convex subset of $\{v \in H(\Omega): v(x_i) \ge 0,\ i=1,\ldots,n\}$ with the property that D contains at least one function which is positive at the data points $x_1,\ldots,x_n$. Then the MPLE of f^o corresponding to penalty function $\phi(v) = \|v\|^2$ in (2.1) exists in D and is unique, where $\|\cdot\|$ denotes the norm on $H(\Omega)$.

Proof: Since $H(\Omega)$ is a RKHS, by the continuity property, for each $i=1,2,\ldots,n$ there exists a constant K_i such that $|v(x_i)| \le K_i\|v\|$. It follows that

$$L(v) = \prod_{i=1}^{n} [v(x_i)]^{d_i} [1-V(x_i)]^{1-d_i} \exp(-\|v\|^2)$$

$$\le \prod_{i=1}^{n} [K_i \|v\|]^{d_i} \exp(-\|v\|^2)$$

$$= \|v\|^k \exp(-\|v\|^2)(\prod_{i=1}^{n} K_i),$$

where $k = \sum_{i=1}^{n} d_i$ is the number of uncensored observations. The function $Q(\lambda) = \lambda^k \exp(-\lambda^2)$, $\lambda > 0$, is bounded above so that $L(v) \le C$, where C is a constant.

Let $M = \sup\{L(v)\colon v \in D\}$. From the hypothesis of the theorem, $M > 0$. There exists a sequence $\{v_j\} \subset D$ such that $L(v_j) \to M$ as $j \to \infty$. Also, since $Q(\lambda) \to 0$ as $\lambda \to \infty$ and $L(v)$ is bounded, $\|v_j\| \le C_1$ for all j, where C_1 is a constant. Now, the set $\{v \in H(\Omega)\colon \|v\| \le C_1\}$ is weakly compact, so $\{v_j\}$ contains a weakly convergent subsequence also denoted by $\{v_j\}$. Let v^* denote the weak limit of $\{v_j\}$. Since $H(\Omega)$ is a RKHS, $v_j(x_i) \to v^*(x_i)$ for each $i=1,2,\ldots,n$. The norm is a continuous convex functional and, hence, is weakly lower semicontinuous. Thus, $\liminf \|v_j\| \ge \|v^*\|$. Since integration is a continuous functional by hypothesis,

$$V_j(x_i) = \int_{-\infty}^{x_i} v_j(t)dt \to \int_{-\infty}^{x_i} v^*(t)dt = V^*(x_i)$$

for all x_i, $i=1,\ldots,n$. Therefore,

$$\lim_{j\to\infty} \prod_{i=1}^{n} [v_j(x_i)]^{d_i} [1-V_j(x_i)]^{1-d_i} \exp(-\|v_j\|^2)$$

$$\le \prod_{i=1}^{n} [v^*(x_i)]^{d_i}[1-V^*(x_i)]^{1-d_i}\exp(-\|v^*\|^2) = L(v^*).$$

Thus, $M \le L(v^*)$. Since D is closed and convex, it is weakly closed, so that $v^* \in D$. Therefore, a maximizer of $L(v)$ exists in D.

Since $M > 0$, we can consider maximizing $J(v)$ over D, where

$$J(v) = \ln L(v)$$
$$= \sum_{i=1}^{n} d_i \ln v(x_i) + \sum_{i=1}^{n} (1-d_i)\ln[1-V(x_i)] - \langle v,v\rangle.$$

The first Fréchet derivative of $J(v)$ is (Tapia, 1971)

$$J'(v)(\eta) = \sum_{i=1}^{n} \frac{d_i \eta(x_i)}{v(x_i)} - \sum_{i=1}^{n} (1-d_i)\eta(x_i) \left[\frac{v(x_i)}{1-V(x_i)}\right] - 2\langle v,\eta\rangle,$$

and the second Fréchet derivative is

$$J''(v)(\eta,\eta) = -\sum_{i=1}^{n} \frac{d_i \eta^2(x_i)}{v^2(x_i)} - \sum_{i=1}^{n} (1-d_i)\eta^2(x_i) \left[\frac{1-V(x_i)+v^2(x_i)}{(1-V(x_i))^2}\right]$$

$$- 2\langle\eta,\eta\rangle.$$

Since $J''(v)$ is negative definite, J is strictly concave by Proposition 16, page 157 of Tapia and Thompson (1978) and, hence, can have at most one maximizer on a convex set. Therefore, there exists a unique solution to (2.1) in D. ///

We note that the constraints in (2.1) define a closed convex subset of $\{v \in H(\Omega): v(x_i) \geq 0,\ i=1,\ldots,n\}$. Also, let (a,b) be a finite interval. For each integer $s \geq 1$, let $H_0^s(a,b)$ denote the Sobolev space of functions on $[a,b]$ whose $s-1$ derivatives are absolutely continuous and vanish at a and b and whose sth derivative is in $L^2(a,b)$. The inner product on $H_0^s(a,b)$ is defined by

$$\langle u,v\rangle = \int_a^b u^{(s)}(t)v^{(s)}(t)dt,$$

where $u^{(s)}$ denotes the sth derivative. It is well known that $H_0^s(a,b)$ is a RKHS with the above inner product and integration over (a,b) is a continuous operation (Lemma 2.1 of De Montricher, Tapia, and Thompson, 1975).

Corollary 3.1. The MPLE corresponding to $H_0^s(a,b)$ with $\phi(v) = \langle v,v\rangle = \|v\|^2$ exists and is unique.

As a special case of Corollary 3.1, we can consider the MPLE of a lifetime density f^o over a finite interval $[0,T]$ for very large $T > 0$ based on an arbitrarily right-censored sample from f^o. The MPLE exists and is unique in $H_0^s(0,T)$ with

penalty function $\phi(v) = \int_0^T [v^{(s)}(t)]^2 dt$. The extension to $[0,\infty)$ is considered in the next section.

4. THE FIRST ESTIMATOR OF GOOD AND GASKINS UNDER CENSORING

For complete samples, Good and Gaskins (1971) considered the penalty function

$$\phi(v) = \alpha \int_{-\infty}^{\infty} \frac{[v'(t)]^2}{v(t)} dt,$$

for $\alpha > 0$, which is equivalent to

$$\phi(v) = 4\alpha \int_{-\infty}^{\infty} \left[\frac{d(v(t))^{\frac{1}{2}}}{dt}\right]^2 dt.$$

De Montricher, Tapia, and Thompson (1975) indicated that the underlying manifold for the MPLE with this penalty function should be $v^{\frac{1}{2}} \in H^1(-\infty,\infty)$, where $H^1(-\infty,\infty)$ is the Sobolev space of functions $f: R \to R$ such that the first derivative f' exists almost everywhere and $f, f' \in L^2(-\infty,\infty)$ with inner product

$$\langle f,g \rangle = \int_{-\infty}^{\infty} f(t)g(t)dt + \int_{-\infty}^{\infty} f'(t)g'(t)dt.$$

Letting $u = v^{\frac{1}{2}}$, we have the penalty function

$$\phi(u^2) = 4\alpha \int_{-\infty}^{\infty} [u'(t)]^2 dt, \quad u \in H^1(-\infty,\infty).$$

This substitution avoids the nonnegativity constraint in problem (2.1).

For the data (x_i, d_i), $i=1,\ldots,n$, described in Section 2, we now would like to maximize

$$L(u) = \prod_{i=1}^{n} [u^2(x_i)]^{d_i} \left[\int_{x_i}^{\infty} u^2(t)dt\right]^{1-d_i} \exp(-4\alpha \|u'\|_2^2).$$

Since $L(u) \geq 0$, maximizing $L(u)$ is equivalent to maximizing $\hat{L}(u) = [L(u)]^{\frac{1}{2}}$. Thus, we have the problem:

$$\text{Maximize } \hat{L}(u) = \prod_{i=1}^{n} [u(x_i)]^{d_i} \left[\int_{x_i}^{\infty} u^2(t)dt\right]^{(1-d_i)/2} \exp(-2\alpha \|u'\|_2^2)$$

$$\text{subject to } \int_{-\infty}^{\infty} u^2(t)dt = 1. \tag{4.1}$$

Letting $J(u) = \ln \hat{L}(u)$, problem (4.1) is equivalent to:

$$\text{Maximize } J(u) = \sum_{i=1}^{n} d_i \ln u(x_i) + \sum_{i=1}^{n} \frac{1}{2}(1-d_i)\ln\left[\int_{x_i}^{\infty} u^2(t)dt\right] - 2\alpha\int_{-\infty}^{\infty}[u'(t)]^2 dt \tag{4.2}$$

$$\text{subject to } \int_{-\infty}^{\infty} u^2(t)dt = 1.$$

Theorem 4.1. Problem (4.2) has a unique solution in the set $S = \{u \in H^1(-\infty,\infty): \int_{-\infty}^{\infty} u^2(t)dt = 1\}$.

Proof: Taking Fréchet derivatives of $J(u)$ gives

$$J'(u)(\eta) = \sum_{i=1}^{n} d_i \frac{\eta(x_i)}{u(x_i)} - \sum_{i=1}^{n} (1-d_i) \frac{\eta(x_i)u(x_i)}{U_2(x_i)} - 4\alpha\int_{-\infty}^{\infty} u'(t)\eta'(t)dt,$$

where $U_2(x_i) = \int_{x_i}^{\infty} u^2(t)dt$, and

$$J''(u)(\eta,\eta) = -\sum_{i=1}^{n} d_i \frac{\eta^2(x_i)}{u(x_i)} - \sum_{i=1}^{n} (1-d_i)\eta^2(x_i) \left[\frac{U_2(x_i) + u^2(x_i)}{(U_2(x_i))^2}\right] - 4\alpha\int_{-\infty}^{\infty} [\eta'(t)]^2 dt.$$

Since $J''(u)(\eta,\eta) < 0$ for $\eta \neq 0$, $J(u)$ is negative definite. Hence, J is strictly concave, and by Theorem 2, page 160, of Tapia and Thompson (1978), $J(u)$ has at most one maximizer in the set

$$S' = \{u \in H^1(-\infty,\infty): \int_{-\infty}^{\infty} u^2(t)dt \le 1\}.$$

If $J(u)$ is continuous on S', by Theorem 4 on page 162 of Tapia and Thompson (1978), J will have at least one maximizer in S'.

Since $H^1(-\infty,\infty)$ is a RKHS, if $u_m \to u$ as $m \to \infty$ in $H'(-\infty,\infty)$, then $u_m(x_i) \to u(x_i)$ for each $i=1,\ldots,n$. Also, $\| u_m - u\| \to 0$ as $m \to \infty$ implies, by definition of the norm in $H^1(-\infty,\infty)$, that $\|u_m - u\|_2 \to 0$ and $\|u_m' - u'\|_2 \to 0$ as $m \to \infty$. Furthermore, for any fixed constant c, $\int_c^{\infty} u_m^2(t)dt \to \int_c^{\infty} u^2(t)dt$ as $m \to \infty$. Hence, $J: S' \to R$ is continuous. Therefore, $J(u)$

has a unique maximizer u_* in S'.

Next, suppose that $\int_{-\infty}^{\infty} u_*^2(t)dt < 1$. Since $u_*^2(t) \to 0$ as $t \to \infty$, then $u_*(t)$ and $u_*'(t)$ both converge to zero as $t \to \infty$. Thus, there exists a number M such that $u_*'(t) < 1$ for $t > M$. Consider a function $v_*(t)$ defined so that (i) $v_*(t) = u_*(t)$ for $t \le M$, (ii) $v_*(t) > u_*(t)$ for $t > M$ and $\int_{-\infty}^{\infty} v_*^2(t)dt = 1$, and (iii) $[v_*'(t)]^2 \le [u_*'(t)]^2$ for $t > M$. Then by (i) and (iii), $\int_{-\infty}^{M}[v_*'(t)]^2 dt = \int_{-\infty}^{M}[u_*'(t)]^2 dt$ and $\int_{M}^{\infty} [u_*'(t)]^2 dt \ge \int_{M}^{\infty}[v_*'(t)]^2 dt$. Also, by (ii), for each x_i, $i=1,\ldots,n$, $\int_{x_i}^{\infty} v_*^2(t)dt > \int_{x_i}^{\infty} u_*^2(t)dt$. These results imply that $J(u_*) < J(v_*)$, a contradiction, since u_* is the unique maximizer of J in S'. Therefore, $\int_{-\infty}^{\infty} u_*^2(t)dt = 1$, completing the proof.

///

Now, we assume that f^o is a lifetime density on the half-line $R_+ = (0,\infty)$ and use a symmetry argument about zero to obtain the results for f^o. Thus, assume that the censored sample (X_i,Δ_i), $i=1,\ldots,n$, is such that $X_i > 0$ with probability one. Then the problem (4.1) becomes:

$$\text{Maximize } \hat{L}(u) = \prod_{i=1}^{n} [u(x_i)]^{d_i}\left[\int_{x_i}^{\infty} u^2(t)dt\right]^{\frac{1}{2}(1-d_i)} \times \exp\left[-2\alpha\int_0^{\infty}(u'(t))^2 dt\right], \tag{4.3}$$

where $x_i > 0$, $i=1,\ldots,n$, $\int_0^{\infty} u^2(t)dt = 1$, and $u(t) \ge 0$, $t > 0$.

Let $X_{-i} = X_i$ and $d_{-i} = d_i$, $i=1,\ldots,n$, and define $\bar{u}(x) = u(|x|)$ for $x \in R\backslash\{0\}$ and $\bar{u}(0) = \lim_{x\to 0^+} u(x)$. Then define the following problem:

$$\text{Maximize } L(\bar{u}) = \prod_{|i|=1}^{n} [\bar{u}(x_i)]^{d_i}\left[\int_{x_i}^{\infty} \bar{u}^2(t)dt\right]^{\frac{1}{2}(1-d_i)} \times \exp\left[-2\alpha\int_{-\infty}^{\infty}(\bar{u}'(t))^2 dt\right], \tag{4.4}$$

where $\int_{-\infty}^{\infty}\bar{u}^2(t)dt = 2$ and $\bar{u} \in H_S \equiv \{g \in H^1(-\infty,\infty): g(x) = g(-x)\}$.

Notice that $L(\bar{u}) = [\hat{L}(u)]^2$. Also, H_S is equivalent to the Sobolev space $H^1(0,\infty)$.

Proposition 4.2. If u^* solves (4.4), then $u_+^*(t) = u^*(t)$, $t \geq 0$, and $u_+^*(t) = 0$, $t < 0$, solves (4.3).

Proof: Suppose u^* solves (4.4). Since $\hat{L}(u) = [L(\bar{u})]^{\frac{1}{2}}$ and u^* is symmetric about zero implies that $\int_0^\infty [u^*(t)]^2 dt = 1$, u_+^* solves (4.3). ///

From Proposition 4.2, the "first MPLE of Good and Gaskins" under arbitrary right-censorship will be given by $(u_+^*)^2(t)$. We next show that this solution exists and is unique.

Theorem 4.3. Problem (4.3) has a unique solution.

Proof: H_S defines a closed convex subset of $H^1(-\infty,\infty)$. Thus, by a proof similar to that of Theorem 4.1, problem (4.4) has a unique solution. By Proposition 4.2, u_+^* is the unique solution to problem (4.3). ///

The next theorem shows that the MPLE from (4.3) has the general form of an exponential spline with knots at the observed data points.

Theorem 4.4. The unique solution u^* of problem (4.4) is an exponential spline with knots at the observed values x_i, $i = \pm 1, \pm 2, \ldots, \pm n$.

Proof: For given $\lambda > 0$ and α in (4.1), let $\phi_\lambda(\bar{u}) = 2\alpha\int_{-\infty}^\infty [\bar{u}'(t)]^2 dt + \lambda\int_{-\infty}^\infty \bar{u}^2(t)dt$ and consider the problem:

$$\text{Maximize} \quad L_\lambda(\bar{u}) = \prod_{|i|=1}^{n} [\bar{u}(x_i)]^{d_i} [\int_{x_i}^\infty \bar{u}^2(t)dt]^{\frac{1}{2}(1-d_i)} \times \exp[-\phi_\lambda(\bar{u})], \tag{4.5}$$

subject to $\bar{u} \in H_S$ and $\int_{-\infty}^\infty \bar{u}^2(t)dt = 2$.

The inner product $\langle u,v\rangle = 2\alpha\int_{-\infty}^\infty u'(t)v'(t)dt + \lambda\int_{-\infty}^\infty u(t)v(t)dt$ defines a norm $\|\bar{u}\|_\lambda^2 = \phi_\lambda(\bar{u})$ equivalent to the original norm on $H^1(-\infty,\infty)$. Let v_i denote the representer in the ϕ_λ-inner product of the continuous linear functional given by point evaluation at x_i, that is $\langle v_i,\eta\rangle_\lambda = \eta(x_i)$ for all $\eta \in H^1(-\infty,\infty)$.

Let $S = \{v \in H_S: v(x_i) \geq 0\}$. Then S is closed and convex. Letting $J_\lambda = \ln L_\lambda$, we have the first and second Fréchet derivatives,

$$J'_\lambda(\bar{u})(\eta) = \sum_{|i|=1}^{n} \frac{d_i \eta(x_i)}{\bar{u}(x_i)} - \sum_{|i|=1}^{n} (1-d_i)\eta(x_i)\left[\frac{\bar{u}(x_i)}{U_2(x_i)}\right] - 2\langle\bar{u},\eta\rangle_\lambda$$

$$= \sum_{|i|=1}^{n} \frac{d_i \langle v_i,\eta\rangle_\lambda}{\bar{u}(x_i)} - \sum_{|i|=1}^{n} (1-d_i) \frac{\bar{u}(x_i)\langle v_i,\eta\rangle_\lambda}{U_2(x_i)} - 2\langle\bar{u},\eta\rangle_\lambda,$$

where $U_2(x_i) \equiv \int_{x_i}^{\infty} \bar{u}^2(t)dt$, and

$$J''_\lambda(\bar{u})(\eta) = -\sum_{|i|=1}^{n} \frac{d_i \eta(x_i)\eta(x_i)}{\bar{u}^2(x_i)}$$

$$- \sum_{|i|=1}^{n} (1-d_i)\eta(x_i)\eta(x_i)\left[\frac{U_2(x_i)+\bar{u}^2(x_i)}{U_2^2(x_i)}\right] - 2\langle\eta,\eta\rangle_\lambda.$$

Thus, due to the nonnegativity of the functions, $-J''_\lambda \geq 2\,\|\eta\|_\lambda^2$ so that $-J''_\lambda$ is uniformly positive definite relative to S. This implies that $-J_\lambda$ is uniformly convex on S. Therefore, if we can show that J_λ is continuous on S, by Theorem 6, page 162 of Tapia and Thompson (1978), J_λ will have a unique maximizer in S.

By an argument similar to that in the proof of Theorem 4.1, if $\bar{u}_m \to \bar{u}$ in H_S as $m \to \infty$, then $\bar{u}_m(x_i) \to \bar{u}(x_i)$ for each x_i and $\|\bar{u}_m - \bar{u}\|_\lambda \to 0$ as $m \to \infty$. Thus $\|\bar{u}_m - \bar{u}\|_2 \to 0$ as $m \to \infty$ and as before, for any fixed constant c, $\int_c^\infty \bar{u}_m^2(t)dt \to \int_c^\infty \bar{u}^2(t)dt$ as $m \to \infty$. In particular, for $c = x_i$, this convergence holds. Therefore, J_λ is continuous on S, and has a unique maximizer $\bar{u}_\lambda$ in S.

Now, at the solution $\bar{u}_\lambda$, we must have the gradient of J_λ vanish, that is,

$$\nabla J_\lambda(\bar{u}_\lambda) = \sum_{|i|=1}^{n} \frac{d_i v_i}{\bar{u}_\lambda(x_i)} - \sum_{|i|=1}^{n} (1-d_i) \frac{\bar{u}_\lambda(x_i)v_i}{U_{2\lambda}(x_i)} - 2\bar{u}_\lambda = 0$$

where $U_{2\lambda}(x_i) = \int_{x_i}^\infty \bar{u}_\lambda^2(t)dt$. Hence,

$$\bar{u}_\lambda = \frac{1}{2}\left[\sum_{|i|=1}^{n} \frac{d_i v_i}{\bar{u}_\lambda(x_i)} - \sum_{|i|=1}^{n} (1-d_i)\frac{\bar{u}_\lambda(x_i)v_i}{U_{2\lambda}(x_i)}\right] . \qquad (4.6)$$

In order to obtain the form of v_i in (4.6), from $\langle v_i,\eta\rangle_\lambda = \eta(x_i)$, we have

$$2\alpha\int_{-\infty}^{\infty} v_i'(t)\eta'(t)dt + \lambda\int_{-\infty}^{\infty} v_i(t)\eta(t)dt = \eta(x_i) . \qquad (4.7)$$

Integrating the left-hand side of (4.7) by parts (in the distribution sense) gives

$$-2\alpha\int_{-\infty}^{\infty}\eta(t)v_i''(t)dt + \lambda\int_{-\infty}^{\infty} v_i(t)\eta(t)dt = \eta(x_i)$$

or

$$\int_{-\infty}^{\infty}[\lambda v_i(t) - 2\alpha v_i''(t)]\eta(t)dt = \int_{-\infty}^{\infty}\delta_i(t)\eta(t)dt, \qquad (4.8)$$

where $\delta_i(t) = \delta_0(t-x_i)$ and δ_0 denotes the Dirac delta function, that is, $\int_{-\infty}^{\infty}\delta_0(t)\eta(t)dt = \eta(0)$. Equation (4.8) is equivalent to the differential equation

$$\lambda v_i(t) - 2\alpha v_i''(t) = \delta_i(t) \qquad (4.9)$$

which, for $i=0$, has the solution

$$v_0(t) = (2\alpha\lambda)^{-\frac{1}{2}} \exp[-(\frac{\lambda}{2\alpha})^{\frac{1}{2}}|t|], \quad t \neq 0.$$

Now, $v_i(t) = v_0(t-x_i) + v_0(t+x_i)$ solves (4.9). Substituting v_i into (4.6) gives the unique solution $\bar{u}_\lambda$ in the form of an exponential spline,

$$\begin{aligned}\bar{u}_\lambda(t) = 2^{-1}(2\alpha\lambda)^{-\frac{1}{2}} \Bigg\{ &\sum_{|i|=1}^{n} \frac{d_i}{\bar{u}_\lambda(x_i)} [\exp(-(\lambda/2\alpha)^{\frac{1}{2}}|t-x_i|) \\ &+ \exp(-(\lambda/2\alpha)^{\frac{1}{2}}|t+x_i|)] \\ &- \sum_{|i|=1}^{n} \frac{(1-d_i)\bar{u}_\lambda(x_i)}{U_{2\lambda}(x_i)} [\exp(-(\lambda/2\alpha)^{\frac{1}{2}}|t-x_i|) \\ &+ \exp(-(\lambda/2\alpha)^{\frac{1}{2}}|t+x_i|)] \Bigg\} . \end{aligned} \qquad (4.10)$$

Now, notice that over the constraints in problem (4.4), problems (4.4) and (4.5) have the same solution for any $\lambda > 0$

since $\int_{-\infty}^{\infty} \bar{u}^2(t)dt$ is constant. We need to show that the unique solution to (4.4), $\bar{u}^*$, is also an exponential spline.

Let $g(\bar{u}) = \int_{-\infty}^{\infty} \bar{u}^2(t)dt$ and

$$G(\bar{u}) = \sum_{|i|=1}^{n} d_i \ln \bar{u}(x_i) + \sum_{|i|=1}^{n} \frac{1}{2}(1-d_i)\ln U_2(x_i) - 2\alpha\int_{-\infty}^{\infty}[\bar{u}'(t)]^2 dt.$$

Then from Lagrange multipliers, there exists λ so that $\bar{u}^*$ satisfies the equation

$$G'(\bar{u}) - \lambda g'(\bar{u}) = 0 \quad \text{and} \quad g(\bar{u}) = 2,$$

$$\sum_{|i|=1}^{n} \frac{d_i \eta(x_i)}{\bar{u}(x_i)} - \sum_{|i|=1}^{n} (1-d_i) \frac{\eta(x_i)\bar{u}(x_i)}{U_2(x_i)} - 4\alpha\int_{-\infty}^{\infty} \bar{u}'(t)\eta'(t)dt - 2\lambda\int_{-\infty}^{\infty} \bar{u}(t)\eta(t) = 0. \tag{4.11}$$

Using L^2 gradients in the sense of distributions, (4.11) is equivalent to

$$\sum_{|i|=1}^{n} \delta_i \left[\frac{d_i}{\bar{u}(x_i)} - \frac{(1-d_i)\bar{u}(x_i)}{U_2(x_i)} \right] + 4\alpha\bar{u}'' - 2\lambda\bar{u} = 0 \quad \text{and} \quad g(\bar{u}) = 2. \tag{4.12}$$

Since (4.4) has a unique solution, (4.12) must also have a unique solution in H_S, namely $\bar{u}^*$.

Now, if $\lambda \le 0$, then any solution of the first equation in (4.12) would necessarily be a sum of trigonometric functions and would not satisfy the constraint $g(\bar{u}) = 2$ since the integral $g(\bar{u})$ would not exist. Thus, $\lambda > 0$. Also $G(\bar{u}) - \lambda g(\bar{u}) = J_\lambda(\bar{u})$, so that $\bar{u}^*$ must solve (4.5) for this λ and therefore has the desired form, an exponential spline with knots at the data points. Hence, the proof is complete. ///

The unique solution to problem (4.3) is then $\bar{u}_+^*(t) = \bar{u}^*(t)$, $t > 0$, from (4.10). Hence, the "first MPLE of Good and Gaskins" is $(\bar{u}_+^*)^2$.

5. CONCLUSION

In this paper we have shown the existence and uniqueness of the MPLE of a density function in an appropriate general mathematical setting, based on arbitrarily right-censored observations from that density. For the first penalty function of Good and Gaskins (1971), the existence and uniqueness of the MPLE of the density function on $(0,\infty)$ was also shown for this type of data. This "first MPLE of Good and Gaskins" under arbitrary right-censoring was shown to be in the form of an exponential spline with knots at the observed censored and uncensored values. These results are analogous to the complete sample case, except that the form of the penalized likelihood, and therefore, the MPLE, is complicated by the terms involving the survival function.

Statistical properties of the MPLE under censoring have not been considered here. The consistency and other statistical results will be investigated in a later paper.

ACKNOWLEDGMENTS

This research was supported by a grant from the United States Air Force Office of Scientific Research and Army Research Office under Grant No. AFOSR-84-0156.

BIBLIOGRAPHY

Barlow, R. E., Bartholomew, D. J., Bremner, J. M., & Brunk, H. D. (1972). Statistical Inference Under Order Restrictions. New York: Wiley.

Bean, S. J. & Tsokos, C. P. (1980). Developments in nonparametric density estimation. Intern. Statist. Rev. 48, 215-235.

De Montricher, G. F., Tapia, R. A. & Thompson, J. R. (1975). Nonparametric maximum likelihood estimation of probability densities by penalty function methods. Ann. Statist. 3, 1329-1348.

Gill, R. D. (1980). Censoring and Stochastic Integrals, Mathematical Centre Tracts 124. Amsterdam: Mathematisch Centrum.

Goffman, C. & Pedrick, G. (1965). A First Course in Functional Analysis. New Jersey: Prentice-Hall.

Good, I. J. & Gaskins, R. A. (1971). Nonparametric roughness penalties for probability densities. Biometrika 58, 255-277.

Klonias, V. K. (1982). Consistency of two nonparametric maximum penalized likelihood estimators of the probability density function. Ann. Statist. 10, 811-824.

McNichols, D. T. & Padgett, W. J. (1982). Maximum likelihood estimation of unimodal and decreasing densities based on arbitrarily right-censored data. Commun. Statist.-Theor. Meth. 11, 2259-2270.

Padgett, W. J. & McNichols, D. T. (1984). Nonparametric density estimation from censored data. Commun. Statist.-Theor. Meth. (Special Issue: Statistical Reviews) 13, (to appear).

Tapia, R. A. (1971). The differentiation and integration of non-linear operators. Nonlinear Functional Analysis and Applications (L. B. Rall, ed.). New York: Academic Press.

Tapia, R. A. & Thompson, J. R. (1978). Nonparametric Probability Density Estimation. Baltimore: The Johns Hopkins University Press.

Wegman, E. J. (1970a). Maximum likelihood estimation of a unimodal density function. Ann. Math. Statist. 41, 457-471.

Wegman, E. J. (1970b). Maximum likelihood estimation of a unimodal density, II. Ann. Math. Statist. 41, 2160-2174

Wertz, W. & Schneider, B. (1979). Statistical density estimation: A bibliography. Intern. Statist. Rev. 47, 155-175.